2020

Expanded Testing of an Integrated qPCR Mixture Screening Assay Using High Resolution Melt Curve Analysis and Support Vector Machine Modeling

Andrea L. Williams
Virginia Commonwealth University

Expanded Testing of an Integrated qPCR Mixture Screening Assay Using High Resolution Melt Curve Analysis and Support Vector Machine Modeling

A thesis submitted in partial fulfillment of the requirements for the degree of Master of Science in Forensic Science at Virginia Commonwealth University.

Andrea L. Williams
Spring 2019-Spring 2020
April 2020

Dawson Cruz Forensic Molecular Biology Research Laboratory
Virginia Commonwealth University
Department of Forensic Science

Tracey Dawson Cruz, Ph.D.
Chair & Professor
VCU Department of Forensic Science

Acknowledgments

I want to extend a great thank you to my committee members Dr. Tracey Dawson Cruz, Dr. Sarah Seashols-Williams, Dr. Edward Boone and Ms. Hannah Wines for entrusting me with continuing this project and their constant support, guidance and assistance. I would also like to thank Mr. Jordan Cox for his assistance with conceptualization, data organization, and trouble shooting issues. Finally, I want to thank the Dawson Cruz laboratory, the Williams laboratory, my family and friends for supporting me through this process.

Abstract

A primary challenge faced by forensic analysts is the demand for timely analysis of evidence. DNA analysis techniques have drastically increased in sensitivity, allowing for low template DNA samples to be detected and used for identification. Since low template samples become even more problematic in a DNA mixture, it would be advantageous to incorporate an assay earlier in the DNA workflow that could detect a mixture and, potentially, determine the number of contributors. Real time PCR instruments have both quantification and high-resolution melt curve analysis (HRM) capabilities allowing for an opportunity to integrate an HRM screening assay into a DNA quantification kit. The melting behavior of DNA varies with nucleotide length/sequence, allowing melt curves to differentiate between single source samples (and their genotypes) and mixtures.

Previously, an assay was developed that integrated an HRM assay (with target STR loci D5S818 and D18S51) into Qiagen's Investigator Quantiplex® kit. Data from this assay was analyzed using linear discriminant and support vector machine (SVM) analyses for sample classification. When the entire HRM curve data was used for classification, genotype prediction accuracies increased to 74.87% for D5 and 26.92% for D18. Further, 100% of mixtures and 87.5% of single source samples were classified as such.

The HRM assay above was evaluated using Quantifiler Trio™ and Investigator Quantiplex® on a more frequently used qPCR platform. When integrated into Quantifiler Trio™, the assay produced inaccurate quantification values, and melt products from the quantification kit were formed. Therefore, Investigator Quantiplex® was used for all subsequent studies, which revealed that genotype prediction accuracies based on STR melt curves were not significantly altered by incorporation into the existing Investigator Quantiplex® kit on the more frequently used qPCR platform. With this integrated Quantiplex/HRM assay, using linear and radial SVM modeling for D5S818 and D18S51 respectively, single source samples (regardless of genotype) and 1:1 mixtures were accurately identified at rates of 42.1% and 60%, respectively. However, these rates dropped when various mixture ratios were tested.

This data confirmed that quantification values and expected melt curves were unaltered, however, overall genotype and mixture predictions reduced and fell below the desired goal (80%). Further, as the minor contributor was reduced, the assay was unable to accurately distinguish between mixtures and single-source samples. Moving forward, it may be necessary to incorporate other mixture ratios into the training set as a way to increase prediction accuracies across a range of mixture ratios.

Key words: Forensic science, qPCR, High resolution melt curves (HRM), genotype, Support vector machine modeling (SVM), Investigator Quantiplex® kit, Quantifiler™ Trio kit, DNA analysis, mixtures, mixture-screening assay

Introduction

The forensic field is continually faced with the challenge of increased, tedious, time-consuming casework along with the criminal justice community's demand for timely analysis of evidence from crime scenes. Short tandem repeat (STR) analysis has long been the primary typing technique used to individualize a sample and create DNA profiles [1]. Over time, forensic analysis methods for DNA amplification and STR profiling have exponentially increased in sensitivity, allowing for low template and/or degraded DNA samples to be detected and to provide profiles useful for identification purposes. However, the interpretation stage of analysis can be less successful with these low quality DNA samples, resulting in outcomes such as allelic drop in and drop out or peak imbalance, making interpretation of the profiles difficult [2]. Low quantity, touch DNA samples can become even more problematic when a DNA mixture is present. Mixtures are samples that include DNA from more than one contributor, and those with low amounts of DNA tend to greatly complicate profile analysis [3]. Unfortunately, in the forensic DNA workflow, the contributor nature of a sample is not revealed until the final step of analysis, during which a mixture is signified by the presence of three of more peaks at multiple loci. Due to these and other potential complications associated with low template mixtures, many laboratories elect not to process or interpret touch DNA samples [4].

If scientists had the ability to detect the contributor nature of a sample earlier on in the DNA workflow, they would have more options for processing and they may, therefore, be less reluctant to process touch DNA samples. For example, if an object is found at a crime scene with anticipated touch DNA evidence, the item would likely be swabbed at several discrete locations (e.g. firearms) using separate swabs in order to collect all possible DNA present while avoiding the possible creation of mixtures. With this situation, it would not be uncommon for each

individual swab to be processed separately, yielding low amounts of DNA, which may be entirely consumed upon initial testing and potentially fail to provide a DNA profile individually. Thus, investigators treat the numerous swabs collected from a given item as individual samples, rather than combining them, in an effort to avoid creating any additional mixtures. However, if an assay were available earlier in the workflow to detect single source samples of a known type, the analyst would be able to decide if it is prudent to combine DNA sources from those individual items (in order to obtain a larger quantity of DNA), prior to STR amplification, in hopes of creating higher quality profiles than if processed individually.

Preliminary work in the Dawson Cruz Laboratory

To address this issue, previous work in the Dawson Cruz laboratory at Virginia Commonwealth University led to the development of an assay to distinguish between single source and mixture samples at the quantification stage of the forensic DNA workflow by integrating a high-resolution melt curve (HRM) screening assay into commercially used quantification kits [5].

Quantifiler™ Trio (Thermo Fisher Scientific, Waltham, MA) and Investigator Quantiplex® (Qiagen, Valencia, CA) are two commercially available DNA quantification kits commonly used by analysts in forensic labs. The Quantifiler™ Trio quantification kit enables forensic laboratories to simultaneously assess the quantity and quality of total amplifiable human female and human male DNA in a sample [6]. This kit uses multi-copy target loci for improved detection sensitivity to assess the level of DNA degradation and/or PCR inhibitors [7]. The human-specific target loci include the small and large autosomal targets, in addition to the male-specific Y target. The primary quantification targets (small autosomal and Y) consist of short

amplicons (approximately 75 to 80 bases) to improve the detection of low template and degraded DNA samples. The Y target is also useful in assessing mixture samples of female and male genomic DNA. The large autosomal target has a longer amplicon (over 200 bases) to assist in determining if a DNA sample is degraded through the use of a ratio of small to long autosomal amplicons [8]. In this multiplex system, the small autosomal target is detected by VIC dye, the large autosomal target is detected by ABY dye and the Y chromosome target is detected by FAM dye [8].

Quantifiler™ Trio uses TaqMan® quantitative real-time PCR chemistry. This strategy employs PCR primers and dye-labeled TaqMan® probes for the amplification of specific loci. During PCR, the TaqMan® probe anneals to a specific complementary sequence between the forward and reverse primer sites. The proximity of attachment between the reporter dye to the quencher results in a suppression of the reporter's fluorescence by Forster-type energy transfer. The cleavage of the probe by the Taq DNA enzyme separates the reporter dye from the quencher, resulting in increased fluorescence by the reporter. This increase in fluorescence, detected by the qPCR instrument, only occurs when the target sequence is complementary to the probe and is amplified during PCR; therefore nonspecific amplification is not detected. [8,9].

The Investigator Quantiplex® quantification kit is another commercially available kit that quantifies human genomic DNA in a sample using quantitative PCR. This kit was designed to confirm if a sample contains enough DNA to allow for downstream analysis and to detect the possible presence of PCR inhibitors. Unlike Quantifiler™ Trio, Investigator Quantiplex® has only one target – the human quantification target. Amplification detection is performed using qPCR scorpion primers chemistry. Scorpion primers are looped, bi-functional molecules that contain a PCR primer linked to a probe and quencher. In its native state, the fluorophore in this

probe interacts with the quencher, reducing fluorescence. During PCR when the probe binds to the PCR products, the fluorophore and the quencher become separated, leading to an increase in fluorescence that is detected by the qPCR instrument. Scorpion primers are known for their rapid hybridization to the target sequence resulting in faster amplicon detection, shorter reaction time and stronger signal, as desired for high throughput assays [10,11].

In the forensic laboratories these quantification kits are used in conjunction with real time PCR or quantitative PCR (qPCR) instruments to detect the amount of amplifiable human DNA in a given sample. The instruments' multi-channel fluorescent detection *and* melt detection capabilities allow for an opportunity to integrate additional assays into existing quantification kits [5]. qPCR instruments amplify, detect and quantify DNA within a sample by combining the function of a thermocycler and fluorometer. Theoretically, fluorescence can be measured during each step of PCR (denaturation, annealing and extension) [12,13]. However, typically, software settings are selected to ensure fluorescence is measured during the extension step of PCR, regardless of the instrument [5]. The Applied Biosystems® 7500 (ABI 7500) (Life Technologies, Carlsbad, CA) is the most frequently used qPCR instrument in the forensic community due to its broad acceptance of producing reliable and accurate data and historical practice within the forensic community [14]. Previous studies in the Dawson Cruz laboratory have used this instrument to detect human genomic DNA and analyze the PCR products through melt curve "dissociation" analysis using fluorescence based PCR reagents [15]. However, the ABI 7500 instrument Sequence Detection Systems (SDS) software has a limited, low resolution dissociation function which generates only 237 data points over the desired temperature range. This low resolution melt capability did not prove to be capable of distinguishing between similarly sized small amplification products [5]; consequently, subsequent work on an integrated

melt curve assay was transitioned to the Rotor-Gene® Q (Qiagen). This qPCR instrument offers quantification, amplification, and high resolution melt curve (HRM) analysis. The Rotor-Gene® Q is equipped with seven dye channels; one being the extended green fluorescent channel (allowing for HRM detection) which generates 1,049 data points over the desired temperature range [16]. Similar to the Rotor-Gene® Q , Life Technologies' QuantStudio 6 (Life Technologies, Carlsbad, CA) is a qPCR platform that has six dye channels while also providing quantification, amplification and high resolution melt curve capabilities, generating 876 data points over the desired range [17]. This instrument is the newest of the three and is currently slowly being introduced to forensic laboratories.

qPCR Melt Curve Analysis

qPCR instruments often utilize special melt curve analysis software, which is used to plot the changes in measured fluorescence of the amplified products over time as the temperature is changed. Fluorescence is typically measured as a result of the signal emitted from an intercalating dye, which as the name implies, intercalates between the nitrogenous bases within double stranded DNA. Thus, as the temperature increases, the DNA denatures (melts), and a corresponding decrease in the fluorescent signal of the intercalating dye is observed. The resulting plot of this data is referred to as a melt curve, from which a derivative plot is generated. The derivative melt curve includes distinct peaks that correspond to the inflection points on the melt curve (**Figure 1)** [5,13,18]. The melting behavior of the dsDNA, and therefore the corresponding melt curve morphology, varies with differences in DNA nucleotide length and sequence, so it may be possible to use these curves to differentiate between various alleles or genotypes of a genetic locus [19]. If specific genotypes can be differentiated using this approach,

it may also be possible to distinguish between mixtures and single source samples and identify them as such [20]. Kuehnert et al. determined that the intercalating dye EvaGreen® (Biotium, Fremont, CA), detected in the green channel at 510nm, could be easily used in the HRM channel of the Rotor-Gene® Q in conjunction with the Investigator Quantiplex® kit, noting that it did not affect the reproducibility of the integrated Quantiplex/HRM assay despite the overlap of emission spectra with the fluorophore in the kit for detection of the human target [20].

Previous studies have utilized STR loci D5S818 and D18S51 for development of an integrated melt curve mixture detection assay [5,21]. D5S818 has small amplicon sizes (115-178 bp), a small range of repeats (6-18), nine known microvariants, and three reported distinct melt curve morphologies [1,22]. Conversely, the STR locus D18S51 has larger amplicons (262-342 bp), a larger range of repeats (7-27), 41 common microvariants, and eight distinct melt curve morphologies [1,22,23]. The difference in amplicon length prevents overlap of resulting D5S818 and D18S51 melt curves, allowing these loci to be easily duplexed (amplified and melted simultaneously in a single reaction) **(Figure 2A)** [20,21].

Further testing of the integrated Quantiplex/HRM assay on the Rotor-Gene® Q revealed that neither the quantification or melt curves produced were altered by the addition of the STR primers and EvaGreen intercalating dye to the Investigator Quantiplex® kit chemistry when HRM thermalcycling parameters were added [5]. The integrated Quantiplex/HRM assay standard curve quality control metrics were all within the acceptable ranges and the melt curves produced distinct, non-overlapping products for both the D5S818 and D18S51 loci without any observed melt products from the Investigator Quantiplex® kit targets [5].

The Rotor-Gene® Q generated 1049 data points per run, which included melt curves for both D5S181 and D18S51, that were exported from the melt curve's software. Kuehnert et al.

characterized melt curves from the D5S181 and D18S51 loci and reported that each had several distinct features, including the heights and temperatures of the primary and shoulder peaks, as well as the ratios between those peaks heights across all samples [20]. Thus, it was suggested that large data sets, exported from the melt curve software for analysis, may, for this purpose, be condensed to a few key elements for statistical evaluation and classification accuracy [20].

Classification Analysis Using Limited Curve Characteristics

To determine if there was a predictable relationship between melt curve morphology and STR genotype, Wines and Cloudy examined melt curves generated after STR amplification using two qPCR instruments (the ABI 7500 and Rotor-Gene® Q). For their initial analysis, they selected only the data that represented the critical characteristic points along the curve ("limited curve characteristics") and they employed several analysis methods for data modeling. The statistical methods evaluated included a commercially available principle component analysis (PCA)-based approach, Rotor-Gene® Q ScreenClust HRM® software, linear discriminant analysis (LDA), and radial and linear support machine (SVM) modeling using R statistical software [5,21]. From previous studies, the PCA-based Rotor-Gene® Q ScreenClust HRM software statistical method was considered limited in that it required the use of multiple standards with every run, which is not ideal for the anticpated use of this integrated assay, and overall, was not accurate enough for classification purposes [20,21]. Therefore, LDA and SVM modeling were taken into consideration and used for future testing. LDA is a classification algorithm that attempts to distinguish variance between and within observations based on other previously classified data patterns [5,24]. Alternatively, SVM modeling's classification algorithm has the ability to apply various statistical algorithms, and model data using large

datasets for classifications based on previously trained observations, similarly to LDA [25,26]. At present, there is not a packaged software available for high-resolution melt curve analysis that utilizes LDA and SVM modeling; however, previous work by the Dawson Cruz laboratory generated code in R statistical software to meet this need [5,21].

In these previous studies, eight to 10 known "training" standard samples of specific genotypes (for both loci) were amplified and melted then used as a basis of comparison for predictive results; in addition unknown samples were tested ("validation" data) including a minimum of 4 samples for each of five to seven genotypes for both STR loci [5,21]. Initial classification attempts using the integrated Quantiplex/HRM assay on the Rotor-Gene® Q and LDA analysis resulted in genotypes being accurately classified for 48.21% and 35.71% of the samples tested for D5S818 and D18S51, respectively. As expected (based on its higher resolution melting capability), the Rotor-Gene® Q out-performed the ABI 7500 in accurate genotyping classification by 34.98% for D5S818 when LDA was used. However, accuracy decreased by 1.02% for D18S51 [5]. In an attempt to improve classification accuracy, training samples with known genotypes that produced similar melt curve morphologies were grouped together in "geno-groupings" for reanalysis [27,28]. As expected, predictions improved when geno-groupings were used, increasing to 64.29% for D5S818 and 62.50% for D18S51 (**Table 1**) [5]. More importantly, when mixtures were incorporated to determine mixture detection accuracies, 56% of the single source samples classified as single source samples (regardless of genotype), and 100% of the mixtures classified as a mixture, for D5S818. For D18S51, 39.39% of the single source samples classified as single source samples (regardless of genotype), and 80% of the mixtures classified as a mixture (**Table 1**) [5].

Classification Analysis Using Whole Curve Data

Although mixture screening using the integrated Quantiplex/HRM assay showed great promise when only key characteristic points from the melt curve were used, accurate prediction of genotypes from single source samples was lower than desired. Utilizing the entire melt curve data available would likely increase the analyst's classification accuracy and power of discrimination. Thus, subsequent studies in the Dawson Cruz lab included prediction modeling using the entire melt curve data set (i.e. "whole curve" analysis). Whole curve analysis allows for data from every temperature point along the curve to be used in the statistical assessment, rather than the (relatively) few characteristic points used in the previous studies. It was expected that the use of all resulting melt curve data could potentially increase the precision and accuracy of resulting genotypic predictions. Thus, whole curve data from single source and mixture samples produced using the integrated Quantiplex/HRM assay on both the Rotor-Gene® Q and the ABI 7500 were analyzed using three statistical models: LDA, SVM with linear basis functions, and SVM with radial basis functions [29]. As with prior analyses, the samples designated as the standards for each genotype were used to train the software prior to analysis of unknown ("validation") samples. As expected, SVM analysis of the entire melt curve data did substantially improve genotyping predictions for both STR loci tested on both qPCR platforms. Also, the Rotor-Gene® Q outperformed the ABI 7500, regardless of classification method for D5S818, and slightly increased the overall genotype accuracy for D18S51 using SVM linear on the Rotor-Gene® Q, in comparison to using the SVM radial on the ABI 7500 (**Table 2)**. Wines' research concluded that SVM modeling, using R statistical software, generated the highest STR genotype prediction accuracies as well as the highest prediction accuracies for determining if a sample was single source or a mixture for both loci tested [5].

When the best performing models for each locus (SVM radial for D5S818 data and SVM linear for D18S51 data) were subsequently used to examine the assay's ability to predict the presence of a mixture, 100% of mixture samples and 87.5% of single source samples were identified as such (**Table 3)**.

Limitations of Integrated Quantiplex/HRM Assay

Although this previously reported work provided an integrated Quantiplex/HRM mixture screening assay that could accurately predict if an unknown sample was a single source or a mixture early in the forensic DNA workflow, there were several limitations of those initial studies that would require further evaluation. Additional work would also be needed to determine the validity, reproducibility and the sensitivity of the assay. Although two person, 1:1 mixtures were correctly identified as such using this assay, the data set analyzed was very small (N=10). Further, samples with different mixture ratios were not evaluated, nor were three, four, or five+ person mixtures tested. Additionally all development and testing of the integrated mixture detection assay previously described was conducted on the higher resolution Rotor-Gene® Q qPCR platform using the Quantiplex quantitation chemistry; the full power of this assay will not be fully realized until the assay is evaluated using a qPCR platform and quantitation chemistry that is more commonly utilized by forensic DNA laboratories.

In an effort to more closely examine the reproducibility, reliability, and robustness of the integrated Quantiplex/HRM mixture-screening assay, reproducibility testing of the integrated assay on the Rotor-Gene® Q was immediately pursued. Further testing of the integrated assay on an updated qPCR platform (QuantStudio 6) was also examined, including the integration of the STR melt assay and subsequent optimization within a more commonly validated human

quantitation kit, the Quantifiler™ Trio. Finally, various two-person mixture ratios (1:2, 1:5, 1:10, 2:1, 5:1, and 10:1) were tested to determine if altering ratios would affect the prediction accuracies and to identify the limit of detection for minor contributors.

Methods & Materials

Over three hundred individuals' buccal swabs were previously collected, extracted, and STR-amplified, and those that expressed a genotype from the chosen list of common D5S818 and D18S51 genotypes were used in this study, as referenced in Cloudy et al[21]. These samples were collected in accordance to the VCU-approved Institutional Review Board (IRB) protocol (HM20006066), which was renewed on December 20, 2019.

Optimized Singleplex HRM Parameters

Prior to initial testing on the ABI 7500 and the Rotor-Gene® Q qPCR platforms, reaction conditions and amplification/melt parameters were optimized for both instruments [5,21]. The "optimized singleplex" reaction conditions for amplification included a 38µl master mix composed of a 1X concentration of Taq Gold Buffer, 3mM $MgCl_2$, 250µM dNTPs, 1µM of each forward and reverse primer [20], 2 units AmpliTaq Gold DNA polymerase (Applied Biosystems), and 2X concentration EvaGreen intercalating dye. Two microliters of sample DNA were added to each reaction for a total reaction volume of 40µl. The "optimized singleplex" amplification cycling parameters were comprised of an initial 10min 95°C denaturation followed by 45 cycles of: 95°C for 5s, 56 °C for 20s, and 65°C for 30s with fluorescence detected during the extension cycle. Following the amplification cycles, samples proceeded through a transition cycle consisting of 72°C for 2min, 95°C for 20s, 55°C for 20s and 56°C for 2min, after which

the amplicons were melted, from 60-95°C at a 0.1° incremental increase, with fluorescent detection in the HRM channel [5,21].

Integrated Quantiplex/HRM Assay on Rotor-Gene® Q

To assess the prior success of integrating the melt curve assay at the quantification step of the DNA workflow, the D5S818 and D18S51 primers [20], and EvaGreen dye were integrated into the commercially available Investigator Quantiplex® quantification kit on the Rotor-Gene® Q, using half volume reactions, as previously described [5]. Integrated Quantiplex/HRM reactions included a 16.16 µl master mix comprised of 7.36µl of the Quantiplex primer mix, 7.36µl of the Quantiplex reaction mix, 100µM forward and reverse primer for each locus, and 0.93X EvaGreen intercalating dye. One microliter of template DNA was added to each reaction for a total reaction volume of 17.16µl. The integrated thermalcycling parameters used ("Quantiplex amplification with transition and melt") varied from the "optimized singleplex" amplification conditions. The integrated assay parameters consisted of a 10min 95°C denaturation followed by 40 cycles of: 95°C for 5s and 60°C for 30s with an added transition stage (72°C for 2min, 95°C for 20s, 55°C for 20s and 56°C for 2 min) between the amplification and melt that occurred (60-95°C at a 0.1° incremental increase) with fluorescent detection in the HRM channel [5,21]. The melt curves were exported for further analysis; each sample's melt data included 1049 data points with melt products ranging from 60°C to 78.49°C for D5S818 and 78.5°C to 95°C for D18S51.

Integration of HRM Assay into Quantifiler™ Trio on QuantStudio 6

Additional testing of the HRM integrated assay continued onto a more commonly used qPCR platform, the Life Technologies' QuantStudio 6, using a more common quantification kit, the Quantifiler™ Trio kit. The reaction conditions' volumes differed, however the amplification/melt parameters were consistent with the "Quantiplex amplification with transition and melt"/integrated parameters in the integrated protocol used on the Rotor-Gene® Q in previous studies [5] and described above.

Unlike the Rotor-Gene® Q, the Quant Studio 6's analysis software did not allow for quantification (amplification) and the high resolution melt to automatically be run consecutively in a single program. Thus, separate quantification and melt protocols were created for the use of the Quantifiler™ Trio on the QuantStudio 6 to achieve the original integrated goal. Prior to assessing the success of the integration of the HRM assay into Quantifiler™ Trio kit, it was necessary to evaluate the possibility of the Quantifiler™ Trio kit producing its own double stranded products in the temperature ranges of interest. In order to do so, the Quantifiler™ Trio kit's standards were amplified using the manufacturer's recommended reaction, but using half reactions, (without STR primers) both with and without the EvaGreen intercalating dye. The conditions for this reaction included a 11µl master mix composed of 4µl of the Quantifiler™ Trio primer mix, 5µl of the Quantifiler™ THP PCR reaction mix, and 0.85X EvaGreen intercalating dye. Two microliters of each standard, used as the template DNA, was added to each reaction for a total reaction volume of 13µl. Amplification and melt parameters were the same as those used in the integrated protocol described above. This generated 876 data points per run, 173 data points fewer than the Rotor-Gene® Q. Additionally, a shift in temperature range

interest for both D5S818 and D18S51 was noted. On this new platform, D5S818 melt products were located between 60-77.48°C and D18S51 products were located between 77.5-95.001°C.

Integrated Investigator Quantiplex/HRM Assay on QuantStudio 6

The reaction conditions and amplification/melt parameters for the integrated Quantiplex/HRM reactions on the QuantStudio 6 were consistent with those described above for the integrated assay on the Rotor-Gene® Q. However, as previously addressed, the quantification and melt occurred separately, back-to-back, on the QuantStudio 6. Amplification and melt parameters were the same as those used in the integrated protocol described above.

In order to assess whether the integrated assay reagents had an effect on the calculated DNA quantities, ten samples were quantified with the standard Quantiplex assay, per the manufacturer's recommended protocol, but using half reactions, as well as with the integrated Quantiplex/HRM assay. Results were compared using a pairwise Student's T-test.

To assess the ability of the integrated Quantiplex/HRM assay on the QuantStudio 6 to distinguish between single source and mixture samples, 20 two-person 1:1 mixtures were created. Mixtures were made based on the previously obtained, known quantities of the single source samples of interest. Samples were diluted to 0.2 ng/µl and equal amounts of DNA from two single source samples were used to create these two-person 1:1 mixtures. In addition, for the assessment of the integrated limit of detection of a minor contributor, mixture ratios (1:2, 1:5, 1:10, 2:1, 5:1 and 10:1) were also created and tested.

In order to determine genotype and mixture detection accuracies, single source and mixture samples available for testing were separated evenly into known "training" and unknown "validation" data sets. For the sake of consistency, the new training set, created on the QuantStudio 6 using the integrated Quantiplex/HRM assay, possessed the same single source

samples and 10 two-person 1:1 mixtures created and used in Wines et al.'s training set created on the Rotor-Gene® Q [5]. This training set consisted of 7-8 samples for each genotype of interest for both D5S818 and D18S51 loci, resulting in a total of 71 samples (41 samples for D5S818, 44 samples for D18S51 and 14 samples overlapping both sets). The training set also included 10 1:1 mixtures, each with genotypes of interest for both loci. In totality, the validation set contained 100 unique single source and mixture samples. The new validation set contained approximately 40 single source samples with genotypes of interest for both D5S818 and D18S51 loci, also used in Wines et al.'s validation set on the Rotor-Gene® Q, plus 10 additional two person 1:1 mixtures, using different single source sample combinations [5].

Classification Analysis

For statistical analysis, the resulting melt curves for both the training and validation sets were analyzed using R software's LDA and SVM (both radial and linear) modeling which allowed for consistency and comparison to Wines et al.'s previous results using the integrated Quantiplex/HRM assay on the Rotor-Gene® Q. For genotyping analysis, prediction accuracies were determined by calculating the percent of single source samples being correctly classified by genotype from the resulting confusion matrices. For mixture screening, prediction accuracies were determined by calculating the percent of single source samples that were classified as such (regardless of genotype) versus the percent of mixtures that were correctly classified as mixtures from the resulting confusion matrices.

To determine the combined accuracy of mixture detection testing, the overall accuracies of both single source prediction and mixture prediction were considered. If a sample classified as a mixture at *either* STR locus, the entire sample was considered a mixture. Conversely, to

classify a sample as single source, it would need to be classified as a single source at *both* STR loci tested (**Table 4**). The combined accuracy was determined by taking the number of correctly classified single source and mixture samples and dividing this number by the total number of samples used within the validation set.

Results

Reproducibility of Rotor-Gene® Q Integrated Quantiplex/HRM Assay

In an effort to examine the reproducibility of the integrated Quantiplex/HRM assay using the Rotor-Gene® Q, a study was conducted that sought to determine the variability of the experimental data from run-to-run. For this, the human DNA quantitation values obtained, the melt curve morphologies, and the prediction accuracies obtained from new runs were compared to those obtained previously [5]. Unfortunately, there were clear differences between the two sets of melt curves generated on the Rotor-Gene® Q (**Figure 2**). Compared to Wines' original melt curve data (**Figure 2A**) [5], D5S818 and D18S51 melt peaks, within the 60-95°C temperature range of interest, were less defined and baselines were higher (**Figure 2B**). Consequently and not surprisingly, melt curves from identical tested samples produced inaccurate genotypes at a much higher rate (data not shown). Interestingly, when the same samples were amplified and melted outside of the Quantiplex kit ("optimized singleplex" reaction), the melt curves were, qualitatively, more similar to Wines' integrated melt curves (**Figure 2C**).

In addition to the high baselines generated in the above mentioned study, the consistent, prominent display of an extra peak between 85 and 87°C was observed. We sought to determine the source of this extra peak and determine if it was a legitimate product of the D18S51's melt morphology, or if it was a product of the Quantiplex kit itself. Thus, the DNA quantitation

standards were used as template DNA in D5S818 and D18S51 amplification and melt reactions using the "optimized singleplex" reaction parameters. For this, individual samples were independently spiked with each component of the Quantiplex kit separately to determine the source of the extra peaks observed. Based on these results, it was concluded that this peak was a result of the STR primers interacting with the DNA in the Quantiplex kit's internal PCR control (IPC) and was not a result of contamination (data not shown). Since this extra peak was determined to be a product of the Investigator Quantiplex® kit, consistently present in every integrated Quantiplex/HRM assay experimental Rotor-Gene® Q run, the extra peak was deemed to be of no consequence to the resulting prediction inaccuracies, which were more likely resulting from the high baseline morphologies observed. In order to rule out instrumental malfunction as the cause of run-to-run inconsistencies, several quality control checks were initiated in order to ensure that the instrument was reaching the appropriate temperatures. Inconsistent temperature readings would be expected to create inconsistent melt curves. Unfortunately, a Rotor-Gene® Q dye calibration cannot be completed outside of the manufacturer's laboratory [16]; however, performing an optical temperature verification test was pursued within the laboratory, as recommended. The temperature verification test passed, however, additional testing continued to reveal inconsistent melt curves. Qiagen's technical support staff suggested that the described issues could be due to a possible temperature shift within the instrument and that having the manufacturer conduct a more precise temperature verification would be the only way to confirm and repair [16]. As a result of this significant expense of this option and the overall limited current use of the Rotor-Gene® Q in forensic laboratories, it was decided to move all further studies to a more frequently used qPCR platform, the QuantStudio 6.

Integration of HRM Assay into Quantifiler™ Trio on Quant Studio 6

In an attempt to integrate the HRM assay into the Quantifiler™ Trio kit, the EvaGreen intercalating dye was added to the quantification kit's normal reaction which was used for amplification and melting of the Quantifiler DNA standards. Prior to using such an assay for mixture screening, it was critical to determine if the added components would affect the expected human DNA quantitation values obtained. In these experiments, the large autosomal target (detected by the ABY custom dye in the yellow channel), the IPC (detected by the JUN custom dye in the orange channel) and the passive reference (detected by the Mustang Purple custom dye in the red channel) were all unaltered. However, the expected quantities in the small autosomal and Y targets were increased substantially due to a commensurate increase in fluorescence caused by the presence of intercalated EvaGreen dye in the amplified kit targets (**Figure 3, Table 5**). EvaGreen's emission wavelength of 530 nm is similar to that of the VIC dye in the blue channel (549 nm, detects the small autosomal target) and of the FAM dye in the green channel (579 nm, detects the Y target). Consequently, it is expected that the EvaGreen dye that is incorporated into these double-stranded products is detected by those channels, inflating the resulting quantitation values. In addition to determining the effects on quantitation, the melt curves of the tested quantitation standards were examined. Unfortunately, the Quantifiler™ Trio kit amplification products did, themselves, create melt products (75-87°C) that overlap with the temperature range expected from the STR integrated assay (60-95.01°C) (**Figure 4**). Due to these interfering factors, in order to integrate this mixture screening assay into the Quantifler™ Trio kit, two significant changes will need to be made. First, a new fluorescent intercalating dye will need to be explored. Ideally, this dye would need to be from a fluorescent detection channel

that does not overlap with those used by the quantitation kit itself (no emission overlap with the VIC, FAM and ABY dyes). Additionally, primers for the loci of interest (D5S818 and D18S51) would need to be redesigned such that the temperature range of their melt products would lie outside of the temperature range of the Quantifiler™ Trio kit melt products.

Integrated Investigator Quantiplex/HRM Assay Testing on Quant Studio 6

In an attempt to transition the integrated Quantiplex/HRM assay to the QuantStudio 6, samples that had been previously quantified using the standard Quantiplex protocol were requantified using the integrated Quantiplex/HRM assay in order to confirm that the added components neither altered the quantitation values obtained nor the expected melt curves produced. In these experiments, concentrations were not significantly different from those previously obtained using the standard protocol, indicating that the addition of the STR primers and the EvaGreen dye did not impact the quantification values obtained ($p = 0.21$, **Table 6**). Further, distinct melt curves for both D5S818 and D18S51 were obtained (**Figure 5**) within the temperature range of interest for the D5S818 and D18S51 products without any additional melt products observed from the kit itself. Ultimately, melt curves produced on the QuantStudio 6 resulted in a lower baseline compared to what was observed in the integrated Quantiplex/HRM assay melt curves on the Rotor-Gene® Q and they were, consequently, deemed adequate for additional testing.

DNA from test ("validation") single-source and 1:1 mixture samples and known standard ("training") samples was amplified and melted using the integrated Quantiplex/HRM assay on the QuantStudio 6. Resulting data from the entire melt curve (whole curve) of the validation samples were compared to the training samples to obtain genotype classifications as well as

single-source vs. mixture classifications using the statistical methods described above based on the confusion matrices produced (**Table 7**). QuantStudio classification accuracies were compared to those obtained previously from the ABI 7500 and Rotor-Gene® Q [5,21] (**Table 8**). The QuantStudio 6 generated an overall lower prediction accuracy for the D5S818 locus and an overall higher prediction accuracy for the D18S51 locus. The highest classification accuracies, on this qPCR instrument, were the same preferred classification techniques found when using the ABI 7500, and were opposite of those identified using the Rotor-Gene® Q. For single source samples, D5S818 genotype prediction accuracies from the QuantStudio 6 were highest when SVM classification methods were used; however, both SVM methods performed the same, producing prediction accuracies of 18.4%. SVM-Radial was the best performing classification method for D18S51 genotype prediction at 31.5%. Using the best classification technique for each locus, 42.1% of single source samples correctly classified as single source (regardless of genotype) and 60% of mixture samples correctly classified as mixture samples (**Table 9**). Both single source and mixture sample prediction accuracies from the QuantStudio 6, as well as overall accuracies, were well above the expected if classified randomly (~6.25%), but were lower than those previously obtained from the Rotor-Gene® Q [5] and less than the desired goal of 80% (**Table 10**).

To assess the ability of the integrated Quantiplex/HRM assay on the QuantStudio 6 to detect minor contributors in a mixture, various two-person mixture ratios were created and tested (including 1:2, 1:5, 1:10 and the inverse of these). Melt curve data was analyzed and samples were classified as either single source or mixture samples as described above using the best performing classification method, as determined above. Unfortunately, at 1:2/2:1, only 5.5% of mixtures were properly identified as such and the overall accuracy of single source vs. mixture

classification was reduced 30.4% (**Table 11**). At 1:5/5:1 (**Table 12**) and 1:10/10:1 (**Table 13**), mixtures were unable to be detected at all, and the overall accuracy rate dropped to 27.6%. Overall, the introduction of mixtures of varying ratios decreased single-source and mixture classification accuracies from 45.8% to 32.3% (**Table 14**).

Conclusions

Previously, an integrated Quantiplex/HRM assay was developed which allowed for accurate prediction of single source vs. mixture DNA samples earlier in the forensic DNA workflow. Having this information early would allow the opportunity for single source samples from the same item to be potentially combined to create higher quality DNA profiles. Unfortunately, this assay was optimized on an qPCR platform that is not common in forensic DNA laboratories (Rotor-Gene® Q) [5,21]. However, previous studies on the most commonly used qPCR platform, the ABI 7500, determined that the resolution capabilities of this platform were not suitable for an STR-based HRM analysis [5]. Consequently, the current work focused on evaluation of the Quantiplex/HRM assay on a newer qPCR platform, the QuantStudio 6 – the newest platform in the ABI qPCR series and newest to be validated and marketed for the forensic DNA community. Additionally, we sought to incorporate the STR-based HRM assay into a quantification kit that is more commonly utilized for forensic casework, the Quantifiler™ Trio.

Unfortunately, when the STR-based HRM assay was integrated into the Quantifiler™ Trio kit, inaccurate quantification values were obtained and the quantitation amplicons targeted in the kit formed a melt product in the range-of-interest for the target STR loci incorporated. Future use of this kit will require the incorporation of a different intercalating dye from an open

dye channel and redesign of STR primers so that the STR melt products are clearly distinguishable from those formed from the quantification targets.

When the original assay (Quantiplex/HRM assay) was initially tested on the QuantStudio 6 qPCR platform, the resulting data confirmed that quantification values and expected melt curves were unaltered, however, overall genotype and mixture prediction accuracies were reduced and fell below the desired 80% when the existing training set data was used. Further, as the minor contributor was reduced (below the 1:1 ratio originally tested), the assay was unable to accurately distinguish mixtures from single sources samples. Going forward, it may be necessary to incorporate other mixture ratios into the training set as a way to increase prediction accuracies across a range of mixture ratios tested. Additionally, although this work exclusively focused on 2-person mixtures, future studies should incorporate additional multi-person mixtures.

While some loss of resolution in the QuantStudio 6 was expected to reduce overall prediction accuracies obtained, the extent of the reduction was unexpected (overall, QuantStudio 6 prediction accuracies were ~43% lower than Rotor-Gene® Q). In the current studies, training samples (standards) were amplified using the integrated Quantiplex/HRM assay, however, previous studies on the Rotor-Gene® Q used training data that was generated using the singleplex chemistry (without the quantitation kit). Thus, future work should repeat the QuantStudio 6 classifications using training data analyzed with the singleplex chemistry. Confidence in the conclusions produced herein will increase when the testing conditions are identical to those previously reported. Fortunately, the best performing SVM methods were identified for QuantStudio 6 melt curve classifications (SVM liner for D5; SVM radial for D18) and this information will prove useful as the project evolves, although it may be wise to evaluate additional classification algorithms, as well.

The work described in this project provides additional proof-of-concept data to justify additional funds to support the continued development of an HRM-based mixture screening assay. Eventually, a simplistic, user-friendly online analysis tool will be developed for easy import of the HRM data for automated genotype prediction and/or single source vs. mixture identification.

Works Cited

[1] J.A. Nicklas, T. Noreault-Conti, E. Buel, Development of a Fast, Simple Profiling Method for Sample Screening Using High Resolution Melting (HRM) of STRs, J. Forensic Sci. 57 (2012) 478–488. doi:10.1111/j.1556-4029.2011.01981.x.

[2] P. Gill, Application of low copy number DNA profiling., Croat. Med. J. 42 (2001) 229–32. doi:10.7213/dialogo.educ.7214.

[3] J. Butler, Forensic DNA Typing, Elsevier Academic Press, New York, 2005.

[4] D.J. Balding, J. Buckleton, Interpreting low template DNA profiles, Forensic Sci. Int. Genet. 4 (2009) 1–10. doi:10.1016/j.fsigen.2009.03.003.

[5] H. Wines, qPCR Genotype Determination and Mixture Detection Using High Resolution Melt Curve Analysis on the Applied Biosystems 7500 and Qiagen Rotor-Gene ® Q qPCR systems, 2018.

[6] C. Vieira-silva, H. Afonso-costa, T. Ribeiro, M.J. Porto, M. Dias, Quantifiler Trio DNA validation and usefulness in casework samples, 5 (2015) 246–247.

[7] N. Gouveia, P. Brito, A. Serra, F. Balsa, L. Andrade, M.S. Bento, P. Cunha, V. Bogas, V. Lopes, M.J. Porto, Validation of Quantifiler1 Trio DNA Quantification kit in forensic samples, Forensic Sci. Int. Genet. 5 (2015) 135–137. doi:10.1016/j.fsigss.2015.09.055.

[8] Applied Biosystems. Quantifiler ™ HP and Trio DNA Quantification Kits, 4485354. (2017) 116. https://assets.thermofisher.com/TFS-Assets/LSG/manuals/4485354.pdf.

[9] Q. Hp, D.N.A. Quantification, K. Cat, Q. Trio, D.N.A. Quantification, K. Cat, Quantifiler ™ HP and Trio DNA Quantification Kits for use with :, (2018).

[10] Qiagen, Investigator® Quantiplex Handbook, (2018).

[11] N. Thelwell, S. Millington, A. Solinas, J. Booth, T. Brown, Mode of action and application of Scorpion primers to mutation detection, Nucleic Acids Res. 28 (n.d.) 3752–3761.

[12] A. Forootan, R. Sjöback, J. Björkman, B. Sjögreen, L. Linz, M. Kubista, Methods to determine limit of detection and limit of quantification in quantitative real-time PCR (qPCR), Biomol. Detect. Quantif. 12 (2017) 1–6. doi:10.1016/j.bdq.2017.04.001.

[13] K.M. Ririe, R.P. Rasmussen, C. Wittwer, Product differentiation by analysis of DNA melting curves during the polymerase chain reaction., Anal Biochem. 245 (1997) 154–60.

[14] A. Kondili, M. Vouropoulou, D. Michos, P. Miniati, Internal validation of the 7500 Real Time PCR System for use in forensic casework in Hellenic Police, Forensic Sci. Int. Genet. Suppl. Ser. 1 (2008) 44–47. doi:10.1016/j.fsigss.2007.12.004.

[15] Applied Biosystems, Applied Biosystems 7000/7500 fast real Time PCR System, (2010). http://tools.thermofisher.com/content/sfs/manuals/4387783c.pdf.

[16] R.Q. User Manual, Rotor-Gene Q Manual, (2012).

[17] QuantStudio™ 6 and 7 Flex Real-Time PCR Systems, Life Technol. (2013). doi:10.1556/OH.2009.28557.

[18] A. Biosystems, Applied Biosystem High Resolution Melting.pdf, (2012).

[19] G.H. Reed, J.O. Kent, C.T. Wittwer, High-resolution DNA melting analysis for simple and efficient molecular diagnostics, Pharmacogenomics. 8 (2007) 597–608. doi:10.2217/14622416.8.6.597.

[20] K. Kuehnert, Development of a Genotype Prediction Screening Assay Using High Resolution Melting Curve Analysis of the STR Loci D5S818 and D18S51, Virginia

Commonwealth University, 2015.

[21] D. Cloudy, Discrimination of D5S818 & D18S51 STR genotypes using LDA & PCA evaluation of high resolution melt curve data from the Qiagen Rotor-Gene Q qPCR Instrument, Virginia Commonwealth University, n.d.

[22] STR Fact Sheet- D5S818, STRBase. (n.d.).

[23] STR Fact Sheet- D18S51, STRBase. (n.d.). http://www.cstl.nist.gov/strbase/str_d18s.htm (accessed February 13, 2019).

[24] S. Balakrishnama, A. Ganapathiraju, Linear Discriminant Analysis - a Brief Tutorial, Compute. (2015).

[25] Y. Tian, Z. Qi, X. Ju, Y. Shi, X. Liu, Nonparallel support vector machines for pattern classification, IEEE Trans. Cybern. 44 (2014) 1067–1079. doi:10.1109/TCYB.2013.2279167.

[26] Rebentrost P, M. Mohsemi, Lloyd S., Quantum support vector machine for big data classification., Phys. Rev. Lett. 113 (2014).

[27] C.N. Gundry, J.G. Vandersteen, G.H. Reed, R.J. Pryor, J. Chen, C.T. Wittwer, Amplicon melting analysis with labeled primers: A closed-tube method for differentiating homozygotes and heterozygotes, Clin. Chem. 49 (2003) 396–406. doi:10.1373/49.3.396.

[28] N. Hong, qPCR Mixture Detection Using Melting Curve Analysis of STR Loci, Virginia Commonwealth University, 2014.

[29] T. Dawson Cruz, Ph.D., Final Draft Summary Dawson Cruz Sept 2018, n.d.

Appendix

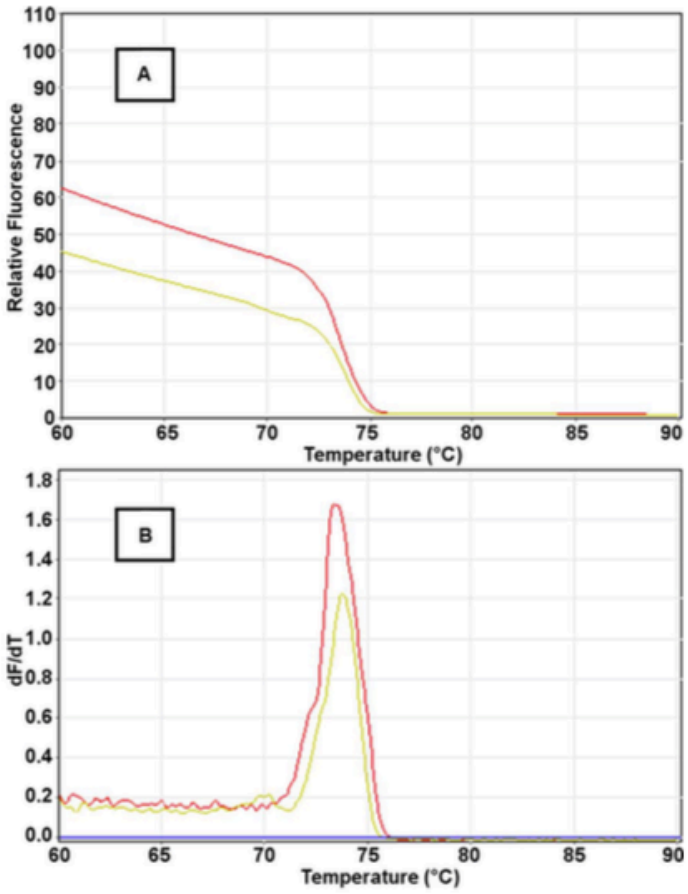

Figure 1. Melt curves and their resulting derivative plots [5]. (A) Curve of melting DNA amplicons measuring fluorescence over time. The inflection point in these samples occur at 73°C (B) Negative derivative plot of melt curves from (A) showing the peaks at the inflection points [5,21].

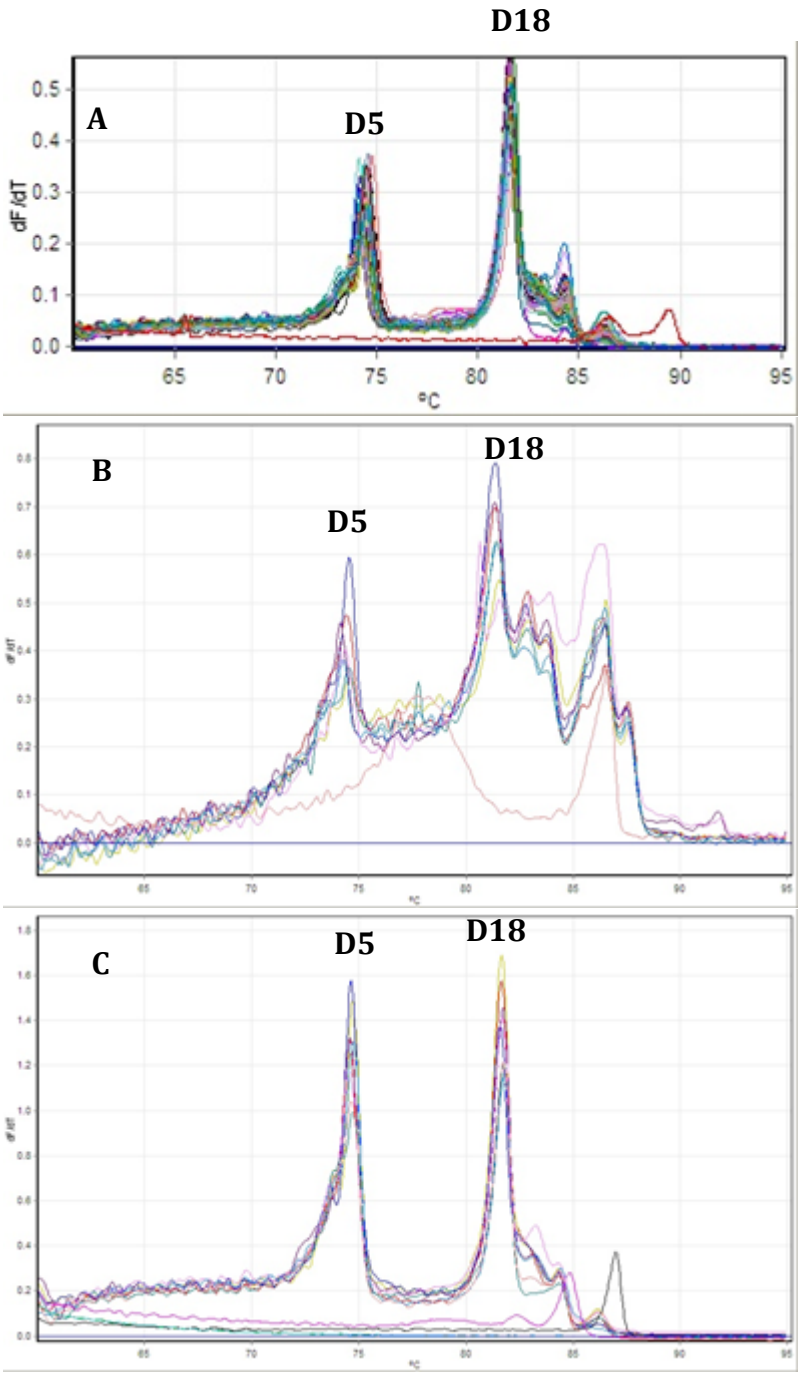

Figure 2: Integrated Quantiplex/HRM Assay on Rotor-Gene® Q melt curves. (A) Melt curves obtained from previously reported studies [5]. (B) Melt curves obtained from reproducibility testing. (C) Melt curves obtained without the quantitation kit, using the "optimized singleplex" protocol.

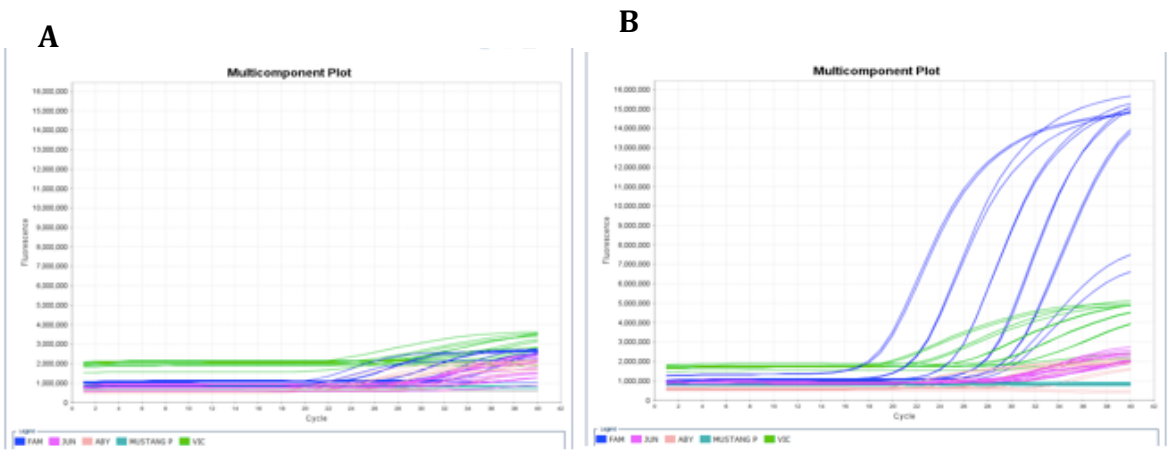

Figure 3: Multicomponent plots from Quantifiler™ Trio kit DNA standards on QuantStudio 6 displaying EvaGreen detection. (A) Samples were run using the standard protocol, with no EvaGreen added. (B) Samples were spiked with EvaGreen, causing a massive increase in fluorescence detected from some of the Quantifiler™ Trio kit's target amplicons in the blue and green channels.

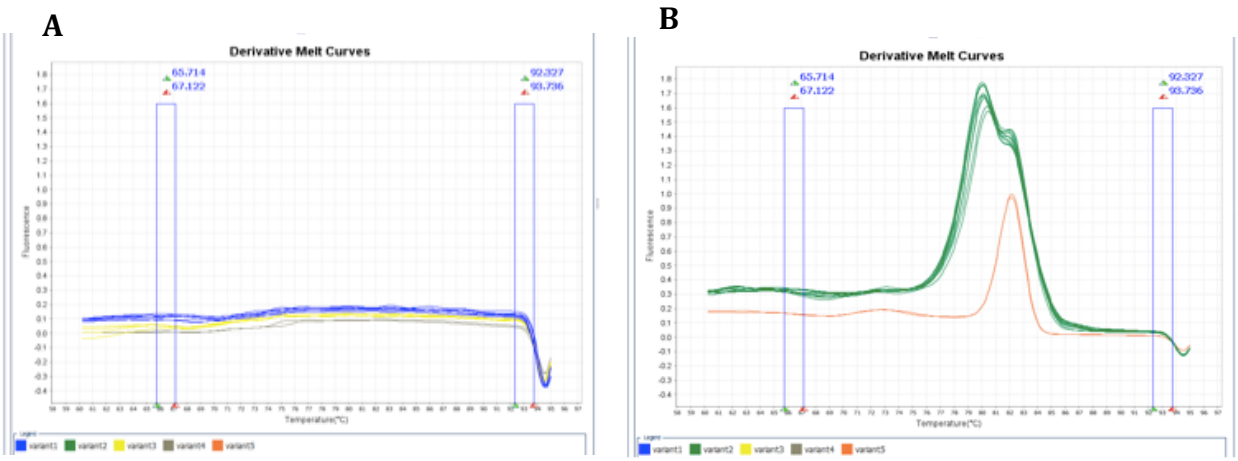

Figure 4: Melt curves from Quantifiler™ Trio kit DNA standards on QuantStudio 6. (A) Samples were run using the standard protocol, with no EvaGreen *or* STR primers added. (B) Samples were spiked with EvaGreen, resulting in a melt product with a primary peak temperature of ~75-87°C.

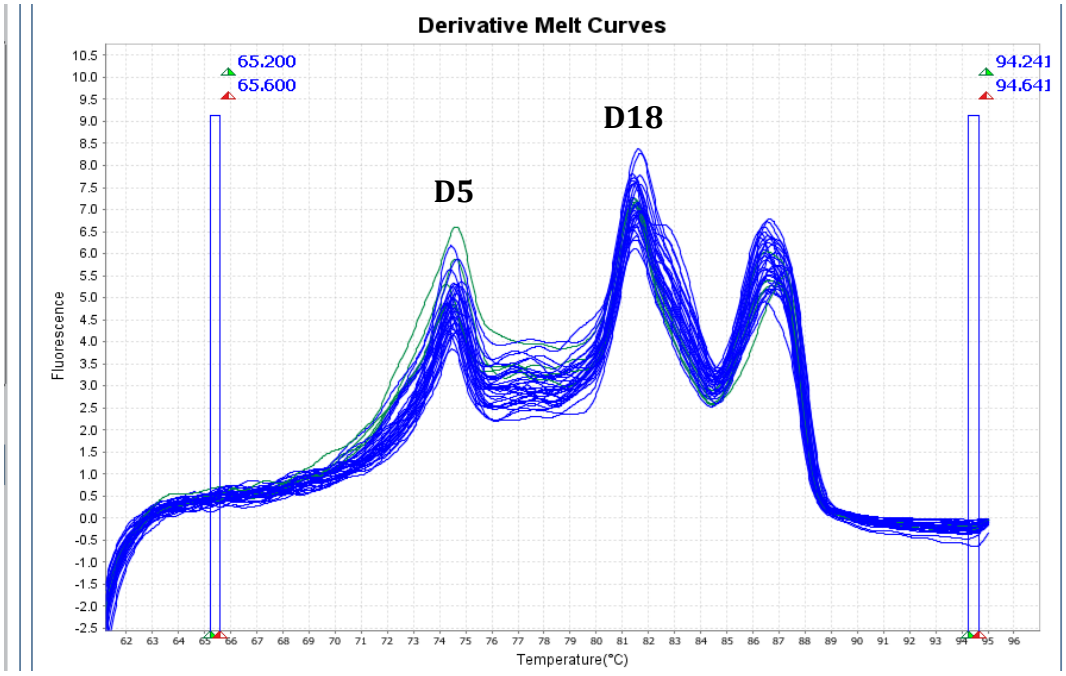

Figure 5. Melt curves from the integrated Investigator Quantiplex®/HRM assay on the QuantStudio 6. Samples were run using the integrated parameters and generated non-overlapping melt curves for D5S181 and D18S51.

Table 1. Previously reported D5S818 and D18S51 classification accuracies using limited melt curve characteristics from the integrated Quantiplex/HRM assay on the Rotor-Gene® Q with LDA [5,21].

	D5S818 (%)			D18S51 (%)		
	Genotype	Geno-group	Mixture	Genotype	Geno-group	Mixture
Single-source n=56	48.21	64.29	N/A	35.71	62.5	N/A
Mixtures + single-source n=26	56	68.18	100	39.39	62.12	80

Table 2. Previously reported D5S818 and D18S51 genotype classification accuracies using whole curve data from the integrated Quantiplex/HRM assay on the Rotor-Gene® Q with three machine learning classification techniques [5,21].

	ABI7500		**Rotor-Gene® Q**	
Technique	D5	D18	D5	D18
LDA	43.39%	9.52%	66.31%	13.46%
SVM-Linear	**50.26%**	16.93%	69.52%	**26.92%**
SVM-Radial	40.21%	**21.16%**	**74.87%**	9.62%

*best method for each locus tested is denoted in bold.

Table 3. Single-source v. mixture prediction accuracies of the integrated Quantiplex/HRM assay for both STR loci tested using whole melt curve data, the Rotor-Gene® Q & the best SVM classification technique [5].

	D5S818 (%)	**D18S51** (%)	**Combined accuracy** (%)
Single-source n=56	94.64	92.86	87.5
Mixtures n=10	100	100	100
		Overall Accuracy (%):	**89.39**

Table 4. Hypothetical data to show how samples classify overall.

Single Source vs. Mixture Prediction			
	D5S818	**D18S51**	**Overall**
Unknown 1	SS	SS	SS
Unknown 2	SS	SS	SS
Unknown 3	M	SS	M
Unknown 4	SS	SS	SS
Unknown 5	SS	M	M
Unknown 6	SS	M	M
Unknown 7	SS	SS	SS
Unknown 8	SS	SS	SS
Unknown 9	M	SS	M
Unknown 10	SS	SS	SS

Table 5: Expected vs. Observed Quantifiler™ Trio Standards without & with EvaGreen dye

			Without EvaGreen (Expected)		With EvaGreen (Observed)	
Sample	**Target**	**Reporter**	**C_t**	**Quantity**	**C_t**	**Quantity**
Std 1a	Small Autosomal	VIC	20.641	50.000	17.708	432.935
	Y	FAM	20.959	50.000	16.286	1,605.090
Std 1b	Small Autosomal	VIC	20.882	50.000	17.994	354.420
	Y	FAM	21.159	50.000	16.454	1,423.884
Std 2a	Small Autosomal	VIC	24.011	5.000	21.007	42.980
	Y	FAM	24.193	5.000	19.345	180.818
Std 2b	Small Autosomal	VIC	24.051	5.000	21.127	39.522
	Y	FAM	24.309	5.000	19.447	168.134
Std 3a	Small Autosomal	VIC	27.452	0.500	24.245	4.452
	Y	FAM	27.838	0.500	22.569	18.115
Std 3b	Small Autosomal	VIC	27.347	0.500	24.370	4.079
	Y	FAM	27.706	0.500	22.619	17.479
Std 4a	Small Autosomal	VIC	30.616	0.050	27.136	0.588
	Y	FAM	30.785	0.050	25.435	2.342
Std 4b	Small Autosomal	VIC	31.089	0.050	27.448	0.472
	Y	FAM	31.597	0.050	25.640	2.024
Std 5a	Small Autosomal	VIC	33.722	0.005	28.994	0.160
	Y	FAM	33.872	0.005	26.911	0.817
Std 5b	Small Autosomal	VIC	33.864	0.005	29.080	0.151
	Y	FAM	33.568	0.005	27.000	0.766

Table 6: Concentrations obtained from single source samples using the Investigator Quantiplex® Assay (left) and the Integrated Quantiplex/HRM Assay (right) on the QuantStudio 6

Sample	Standard Quantiplex Assay (ng/µl)	Integrated Quantiplex/HRM Assay (ng/µl)	Percent Difference
2224	9.344	6.020	35.57%
2235	7.285	4.976	31.70%
2259	2.487	0.483	80.58%
2269	0.63	0.598	5.08%
2292	4.116	6.21	-50.87%
2299	2.913	4.872	-67.25%
2329	23.835	5.338	77.60%
2331	5.549	2.052	63.02%
		Average Difference (%):	21.93 ± 56.12

$p = 0.21$

Table 7. Confusion matrices generated using R software and best SVM classification technique for both D5S181 and D18S51. (A) Confusion matrices generated for genotype prediction accuracies (B) Confusion matrices generated for genotype, single source and mixture prediction accuracies

A

Genotype Prediction - Confusion Matrices

D5S818

Predicted

Actual \ SVM Linear	1011	1111	1112	1113	1212	1213	1313
1011	**1**	0	1	0	0	0	1
1111	2	**2**	0	0	0	0	0
1112	9	2	**1**	1	0	1	2
1113	1	1	1	**1**	0	0	0
1212	1	0	0	0	**1**	0	0
1213	1	1	2	1	0	**0**	0
1313	1	0	1	1	0	0	**1**

Single Source Genotype Accuracy: 18.4%

D18S51

Predicted

Actual \ SVM Radial	1213	1214	1215	1216	1314	1316	1415
1213	**2**	0	1	0	2	1	0
1214	1	**2**	0	0	2	0	0
1215	0	2	**1**	0	2	0	0
1216	0	0	0	**1**	1	2	0
1314	0	3	1	1	**2**	2	1
1316	0	0	0	0	2	**3**	0
1415	0	0	0	1	0	1	**1**

Single Source Genotype Accuracy: 31.5%

B

Single Source vs. Mixture Prediction - Confusion Matrices

D5S818

Predicted

Actual \ SVM_Linear	1011	1111	1112	1113	1212	1213	1313	mix
1011	**0**	0	0	0	0	0	1	2
1111	0	**1**	0	0	0	0	0	3
1112	2	0	**1**	1	0	1	2	9
1113	0	0	0	**1**	0	0	0	3
1212	0	0	0	0	**1**	0	0	1
1213	0	1	2	1	0	**0**	1	0
1313	0	0	0	1	0	0	**0**	3
mix	0	4	0	0	0	0	0	**6**

Single Source Accuracy: 44.7%
Mixture Accuracy: 60%

D18S51

Predicted

Actual \ SVM_Radial	1213	1214	1215	1216	1314	1316	1415	mix
1213	**1**	1	1	0	1	0	0	2
1214	1	**2**	0	0	1	0	1	0
1215	0	2	**1**	0	2	0	0	0
1216	0	0	0	**1**	1	2	0	0
1314	0	2	1	1	**2**	1	2	1
1316	0	0	0	0	2	**2**	0	1
1415	0	0	0	1	0	1	**1**	0
mix	0	10	0	0	0	0	0	**0**

Single Source Accuracy: 89.5%
Mixture Accuracy: 0%

Table 8. Classification accuracies for whole curve analysis for ABI 7500, Rotor-Gene® Q and QuantStudio 6 for classification techniques: LDA, SVM-Linear and SVM-Radial.

	ABI 7500		Rotor-Gene® Q		QuantStudio 6	
Technique	**D5**	**D18**	**D5**	**D18**	**D5**	**D18**
LDA	43.39%	9.52%	66.31%	13.46%	7.90%	23.60%
SVM-Linear	**50.26%**	16.93%	69.52%	**26.92%**	**18.40%**	10.50%
SVM - Radial	40.21%	**21.16%**	**74.87%**	9.62%	18.40%	**31.50%**

*best method for each locus tested is denoted in bold.

Table 9. Single source v. 1:1 mixture prediction accuracies of the integrated Quantiplex/HRM assay using whole melt curve data, the QuantStudio 6 & the best SVM classification technique.

	D5S818	**D18S51**	**Combined Accuracy**
Single-source N=38	44.7%	89.5%	42.1%
Mixtures N=10	60%	0%	60%
		Overall accuracy	45.8%

Table 10. Single source v. 1:1 mixture prediction accuracies of the integrated Quantiplex/HRM assay using whole melt curve data, the Rotor-Gene® Q, QuantStudio 6 & the best SVM classification technique.

	Combined Accuracy	
	Rotor-Gene® Q	**QuantStudio 6**
Single-source N=38	87.5%	41.2%
Mixtures N=10	100%	60%
Overall accuracy	**89.39%**	**45.8%**

Table 11. Single source v. 1:2/2:1 mixture prediction accuracies of the integrated Quantiplex/HRM assay using whole melt curve data, the QuantStudio 6 & the best SVM classification technique.

	D5S818	**D18S51**	**Combined Accuracy**
Single-source N=38	44.7%	89.5%	42.1%
Mixtures N=18	5.5%	0%	5.5%
		Overall accuracy	**30.4%**

Table 12. Single source v. 1:5/5:1 mixture prediction accuracies of the integrated Quantiplex/HRM assay using whole melt curve data, the QuantStudio 6 & the best SVM classification technique.

	D5S818	**D18S51**	**Combined Accuracy**
Single-source N=38	44.7%	89.5%	42.1%
Mixtures N=20	0%	0%	0%
		Overall accuracy	**27.6%**

Table 13. Single source v. 1:10/10:1 mixture prediction accuracies of the integrated Quantiplex/HRM assay using whole melt curve data, the QuantStudio 6 & the best SVM classification technique.

	D5S818	**D18S51**	**Combined Accuracy**
Single-source N=38	44.7%	89.5%	42.1%
Mixtures N=20	0%	0%	0%
		Overall accuracy	**27.6%**

Table 14. Combined Single source & mixture ratio prediction accuracy of the integrated Quantiplex/HRM assay using whole melt curve data, the QuantStudio 6 and best SVM classification technique

	1:1 Mix	**1:2 & 2:1 Mix**	**1:5 & 5:1 Mix**	**1:10 & 10:1 Mix**	**Overall Accuracy**
Single–source N=38	42.1%	42.1%	42.1%	42.1%	42.1% N=152
Mixtures	60% N=10	5.5% N=18	0% N=20	0% N=20	10.3% N=68
Combined Accuracy	**45.8%**	**30.4%**	**27.6%**	**27.6%**	**32.3%**

Vita

Andrea (Dre) Lynne Williams is a Plainfield, New Jersey native; who's love for forensic science guided her to Virginia Commonwealth University (VCU) in Richmond, VA. She received her Bachelor's of Science in Forensic Science with a concentration in Forensic Biology, dual minoring in biology and chemistry, with Cum Laude honors from VCU in May 2018. She is currently pursuing her Master's of Science in Forensic Science with a concentration in Forensic Biology, anticipating graduation in May 2020 with Cum Laude honors from VCU. Since beginning her journey within the graduate program, Andrea has worked as a Graduate Research Assistant and Quality Assurance/Quality Control manager within the Dawson Cruz Forensic Molecular Biology laboratory, located in VCU's Department of Forensic Science. In addition to taking on these new roles, Andrea has had the opportunity to attend the 71st Annual American Academy of Forensic Sciences Conference in February 2019 as a student affiliate and has presented her directed research at an international forensic DNA conference, the 30th Annual International Symposium on Human Identification (ISHI), in September 2019. As for academic accomplishments, Andrea has made Dean's list, was awarded the Emily R. Murphy scholarship in February 2019 and the "Black History in the Making Award" by VCU's Department of Forensic Science in February 2020.